BBC RD 1974/38

THE QUEEN'S AWARD TO INDUSTRY 1969

THE QUEEN'S AWARD TO INDUSTRY 1974

RESEARCH DEPARTMENT REPORT

Properties of hearing related to quadraphonic reproduction

P.A. Ratliff, B.Sc., Ph.D.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION November 1974

BBC RD 1974/38
UDC 534.76

PROPERTIES OF HEARING RELATED TO QUADRAPHONIC REPRODUCTION

P.A. Ratliff, B.Sc., Ph.D.

Summary

An investigation of some of the properties of hearing relevant to the reproduction of an omnidirectional sound-stage has been undertaken. The work was designed to provide basic knowledge of certain properties of hearing, and to examine under critical conditions some of the phenomena which occur with quadraphonic (four-loudspeaker) reproduction of a sound field.

Three fundamental, sound-locating properties of the human auditory system have been determined, and a law has been established relating interchannel level and image location around the listener, using quadraphonic sound-source arrangement. Effects of unwanted signals in the reproduced field are examined, and also the effects of phase-shifts inserted between these signals, such as those which typically occur in the matrix quadraphonic systems currently under consideration by many workers. Results expose some psycho-acoustic myths, and account for some of the observed peculiar phenomena in practical matrix systems.

Issued under the authority of

P. Rainger

Head of Research Department

Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION
November 1974

(PH-129)

PROPERTIES OF HEARING RELATED TO QUADRAPHONIC REPRODUCTION

PROPERTIES OF HEARING RELATED TO QUADRAPHONIC REPRODUCTION

P.A. Ratliff, B.Sc., Ph.D.

Terminology

For brevity in this report abbreviations for directions with respect to the listener are used extensively, along with a few other commonly used terms which are listed below.

A linear sixteen point direction scale for the horizontal plane is introduced in which direction '0' is always directly in front of the listener. Even numbered directions are also referred to by an alphabetic code indicating the directions in words, as shown in Diagram A.

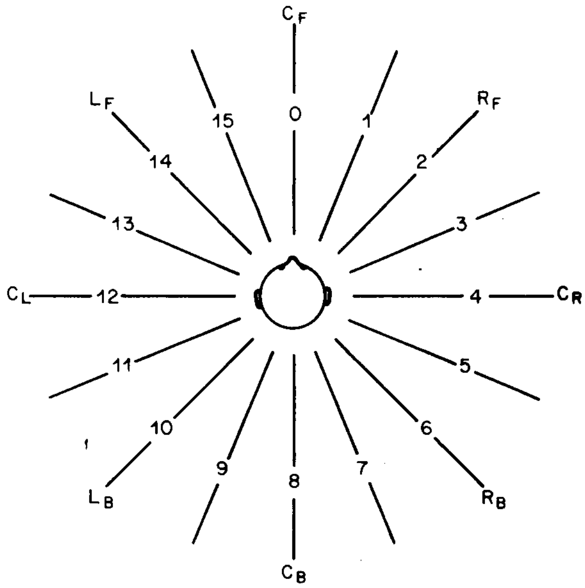


Diagram A

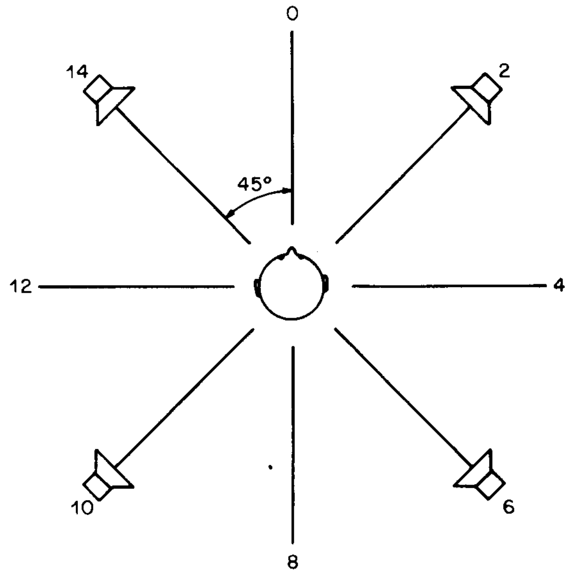
C_F centre-front, R_F right-front, C_R centre-right, R_B right-back
C_B centre-back, L_B left-back, C_L centre-left, L_F left-front

There are two well known quadraphonic arrangements of loudspeakers, the 'square' and the 'diamond' arrays as shown in Diagram B.

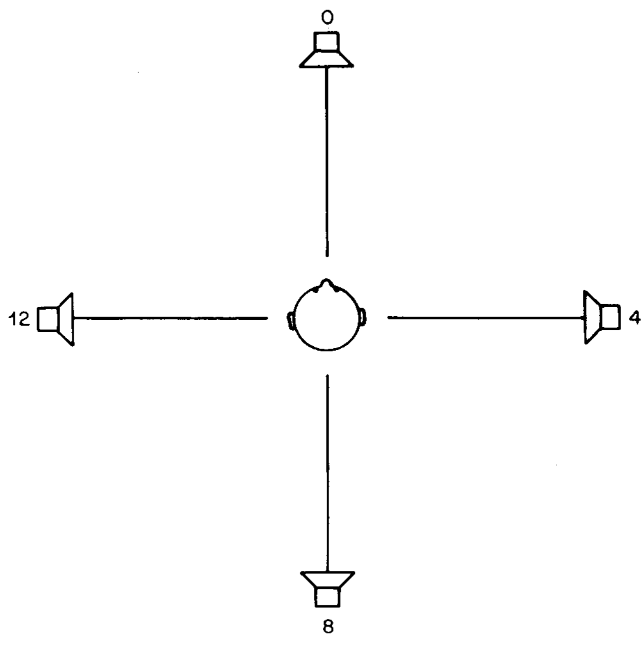
Much of this report deals with the 'square' array in which the loudspeaker positions are sometimes referred to as 'corner locations', and the span between any adjacent pair of loudspeakers is referred to as a 'quadrant', specifically defined as 'front' (14–2), 'right' (2–6), 'back' (6–10) or 'left' (10–14). Directions 0, 4, 8 and 12 are then referred to as centre-quadrant directions.

Other abbreviations used are listed below:

l. s.	loudspeaker
m.p.*l*.	minimum perceptible level
s.d.	standard deviation
s.p.l.	sound pressure level



Square Array



Diamond Array

Diagram B

The following sound quality abbreviations are used:

B bass heavy
C close, near
D diffuse
F far, distant
G good, no adverse comment
H high)
L low) vertical position
NH normal)
J jumpy, horizontal image location jumps between loudspeakers
N nasal, bass lacking
P phasey, in the head sensation
s slightly (used in conjunction with above, i.e. sD, slightly diffuse)
v very (used in conjunction with above, i.e. vD, very diffuse).

1. Introduction

Increasing demands for standards on quadraphonic 'surround-sound' reproduction have led to the realisation that a greater fundamental knowledge of the properties of the human auditory system is required, if an optimum technical solution is to be obtained. This report contains experimental results on the angular (azimuthal) localisation and subjective quality of real and imaginary sound sources, and considers the effects of 'unwanted signals' in four-loudspeaker reproductions of uniquely localised images, typical of those produced by some recording techniques, and by matrix quadraphonic systems.

An understanding of the processes of human hearing is at present far from complete, and although there is a considerable quantity of literature on the subject (e.g. see references of Ref. 1), theories presently proposed[2,3] do not account satisfactorily for all observed phenomena. However, for present requirements, it is necessary to determine the fidelity with which a quadraphonic system can reproduce a sound field which *subjectively* satisfies the listener.

2. Equal loudness levels

It is well known that the sound levels at each ear differ, particularly at high frequencies, depending on the azimuthal direction of the source,[2] and this has been put forward to support an inter-aural intensity hypothesis of localisation.[4] It is thought unlikely, however, that this frequency-dependent difference is the sole determining factor. There is greater support for the inter-aural time-difference hypothesis,[5] although this is probably an over simplification of the localising process.[3]

An experiment to determine the subjectively-assessed equal-loudness levels around an observer was conducted,* in which the observer was presented with sound from one

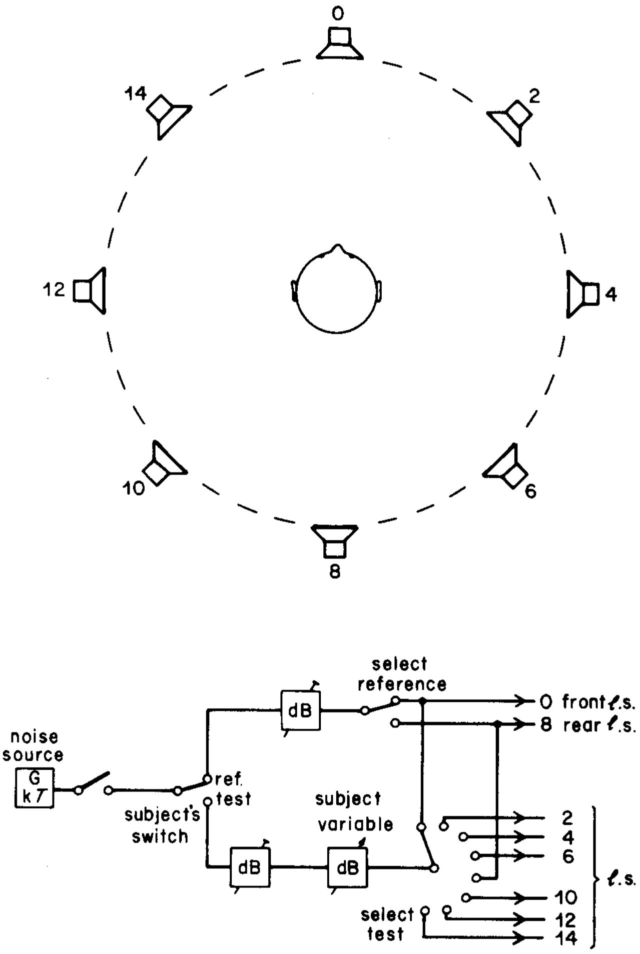


Fig. 1 - Arrangement for 'equal-loudness levels' experiment

of eight closely-matched loudspeakers arranged symmetrically around him. He was provided with a switch and a calibrated attenuator, and asked to adjust the attenuator until the sound-source was of the same loudness as a fixed reference, which could be selected by operating the switch. The reference loudspeaker was always immediately in front of, or behind the subject, who was seated in a chair with the nape of his neck against a thin wooden head-rest. The subject was asked to keep his head still, facing the front throughout the test; the arrangement is shown in Fig. 1.

The experiment was conducted in a specially designed 'average' listening room (about 70 m^3 capacity and average reverberation time of 0·35 sec.) using octave bands of pink-noise, centred on 230 Hz, 2 kHz and 7 kHz, reproduced at normal listening levels (about 70 dBA) from eight high-quality loudspeakers equally spaced on a circle 3·35 m in diameter. The results of tests using eight observers are shown in Fig. 2(*a*), and show little variation with azimuth. The experiment was repeated in a free-field room (surface reflection less than 10% at all frequencies above 40 Hz), but realistic results could only be obtained in the 2 kHz band, because the degree of loudspeaker matching was insufficient for observers not to be disturbed by subjective quality differences between them at other frequencies; the degree of subjective quality matching required under non-reverberant conditions such than an observer can make a loudness assessment with conviction is very high indeed,

* Work undertaken by T.W.J. Crompton.

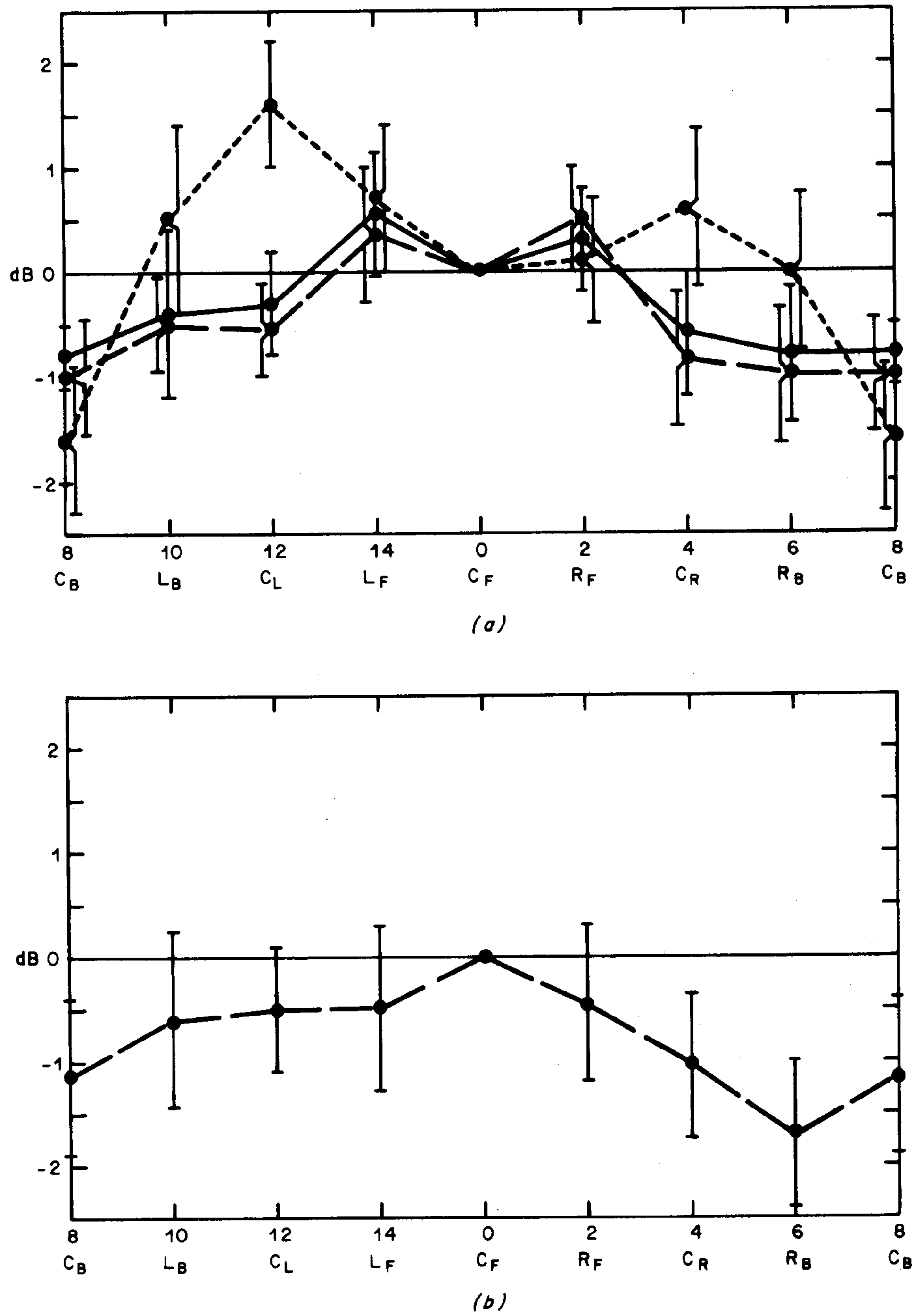


Fig. 2 - Equal-loudness levels of octave bands of pink noise

(*a*) Listening room (*b*) Free-field room

——— 230 Hz) band
— — 2 kHz) band
- - - - 7 kHz) band

experimental result showing standard deviation

but small mis-matches are effectively masked by the reverberant fields present in typical listening rooms. The result of the free-field room experiment is plotted in Fig. 2(*b*), and again shows little variation with azimuth. It is concluded that the average auditory response is almost equally sensitive around the full azimuth circle, although there is a consistent trend for the back to be less sensitive than the front by about 1 dB, increasing slightly at high frequencies.

3. Absolute perception of direction

An important feature in 'surround-sound' reproduction systems is their ability to reproduce the directional properties of the programme to a subjectively acceptable standard,

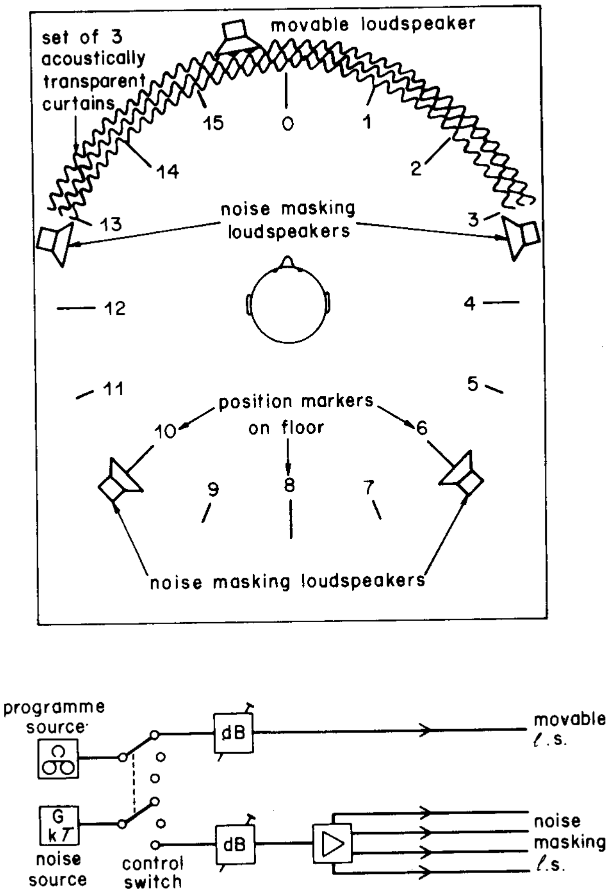


Fig. 3 - Listening room arrangement for localisation experiment

and to this end the absolute azimuthal accuracy of the auditory system was determined.* Initial tests were performed in the listening room with the arrangement shown in Fig. 3. The subject faced acoustically transparent curtains or sat at some multiple of 90° to this direction to examine each quadrant separately. This was necessary as there was insufficient room to completely surround the subject by curtains and examine the full compass in one experiment. The subject was asked to locate the sound-source (a moveable loudspeaker) in each of several tests, which were interposed with masking noise (from the four loudspeakers outside the curtains) to conceal any loudspeaker movement noises. The subject was given a chart (see Fig. 4) dividing the compass into sixteen 22½° segments (units), and was initially asked to place the sound-source in relation to this scale; markers were also placed on the floor to assist in angular awareness. The programme material for these tests consisted of a repeated 30 second excerpt of percussive music. This was found, in preliminary work, to be the most critical material for image localisation assessments.

* 'Absolute azimuthal perception' is defined as that related to the localisation of a single sound-source, and the term 'relative azimuthal perception' is used to denote that describing the relative localisation of one sound-source with reference to another that is closely spaced to it.

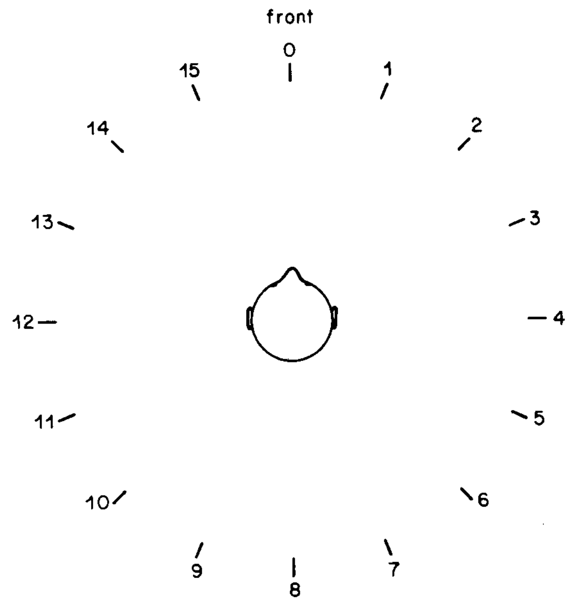


Fig. 4 - Location chart used in localisation experiments

A number of psychological problems were encountered in finding a suitable locating method. Quantisation of the position assessment occurred whether the sound-source was placed on a marker or between them, and since an increase in the number of marker positions would merely serve to confuse the subject, a second method was used where the experimenter moved the loudspeaker until the subject was satisfied that the source was where he had been instructed to locate it. During all such movement the loudspeaker was switched off and masking noise was switched on to avoid giving localising information by movement, and to dissociate each test. This second method proved more acceptable, and results of seven observers are shown in Fig. 5.

The front-stage quadrant is found to be fairly accurately defined (standard deviation, s.d. $\triangleq \pm 2{\cdot}5^\circ$) with marginal image expansion at the extremes of the quadrant. Centre-back (C_B) is similarly well defined, but away from this location greater uncertainty of source position arises. Greatest uncertainty occurs at left-back (L_B) and right-back (R_B), and considerable rear-image expansion occurs; about 11° at L_B and R_B. The amount of image-shift and the standard deviation are greater in rear-quadrant examination than in side-quadrant examination, for the same nominal source positions.

This is thought to be a psychological phenomenon, probably due to visual cues which can modify the ability of the brain to make unbiased decisions. For instance, the finite width of the curtains and the knowledge that the loudspeaker was always constrained to be located behind the curtains could have given rise to the discrepancies observed during the tests.

These results, and comments made by the subjects, indicate a considerably greater ease in the localisation of sound-sources which appear in the front 'visual' quadrant. The ability to 'see' the sound source (although it was, of course, concealed behind the curtains) appears to improve

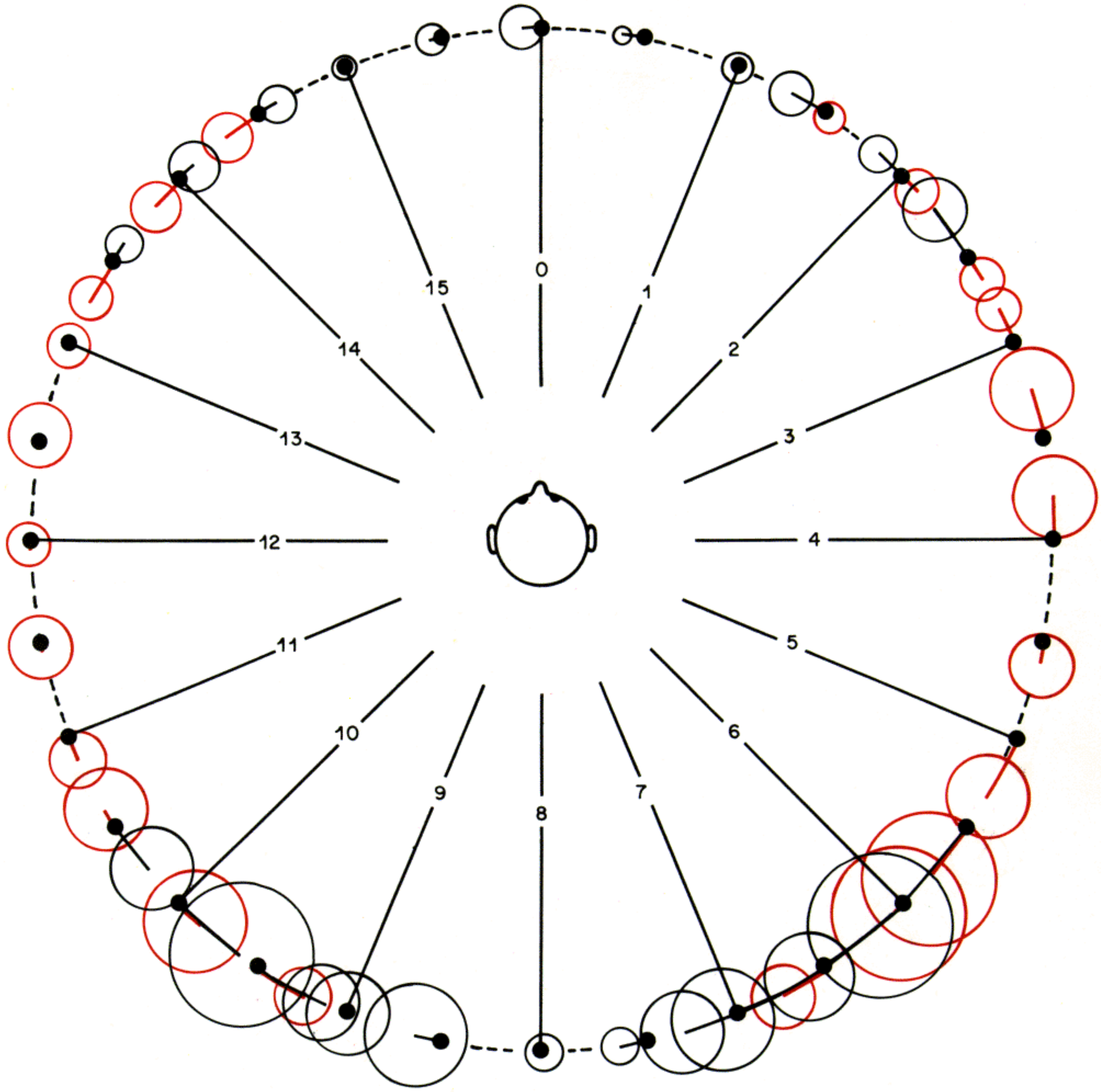


Fig. 5 - Absolute sound localisation in the listening room

* The circles denote the measured standard deviation (s.d.) of the source position. Its radius subtends the angular s.d. as perceived by the subject.

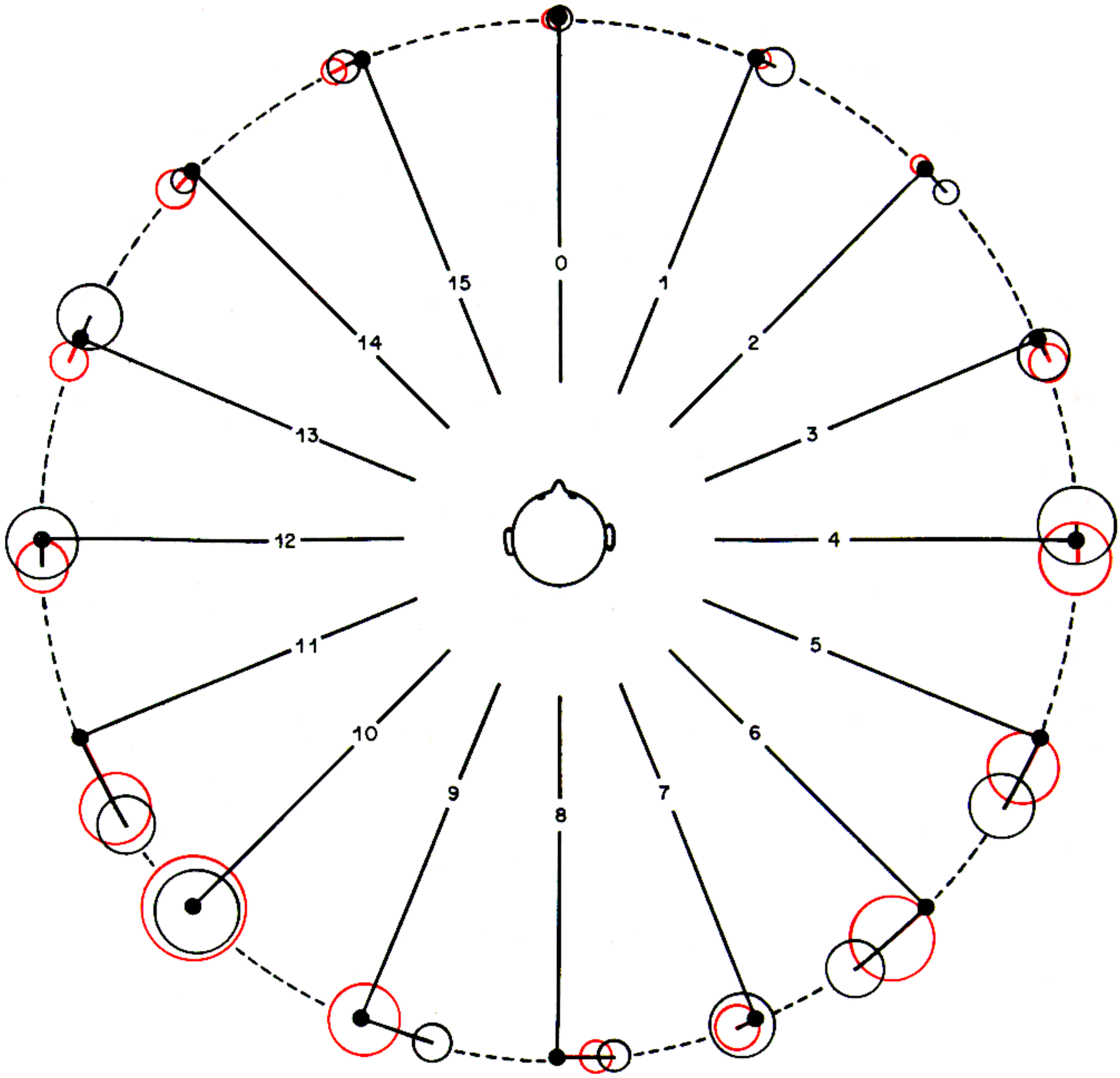
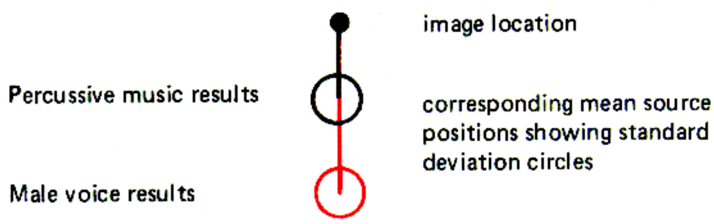


Fig. 6 - Absolute sound localisation in the free-field room



the subjects' ease of sound-source localisation. This is possibly because the brain readily correlates information from multiple sensory inputs. In order to remove the effect of visual information, five subjects repeated the experiment in the front quadrant, blindfold, and their results showed remarkable similarities to those obtained for the rear quadrant previously. The uncertainty of position became similar to that in the rear quadrant, and image expansion again occurred, although only at the extremes of the front quadrant.

The anomalies observed made further use of the listening room undesirable, and so further work was conducted in the free-field room. There was sufficient room to construct a complete circle of curtains, 2 m in diameter, at the centre of which the subject was seated, with his head against the head-rest, facing centre-front (C_F). Similar position markers were affixed around the curtains, and the sound-source (a compact high-quality loudspeaker unit) was moved around a 1·5 m radius circle, beyond the curtains, at a height just below the subject's ears. Four loudspeakers in the corners of the free field room provided movement masking noise, as in the listening room. Illumination was provided only within the curtains to ensure that they formed a visually opaque screen to the subject. Results for seven subjects using percussive music are shown in Fig. 6 (in black), and are substantially more consistent than those obtained in the listening room.

Azimuthal acuity in the front semi-circle is good with little error, s.d.s only reaching about $\pm 3{\cdot}5^\circ$ at the extremes (C_L and C_R). Again acuity at C_B is equal to that at C_F (s.d. $\triangleq \pm 1^\circ$), but a small left-hand image offset is observed. This is thought to be due to the entrance into the curtains being right of C_B, which may have pre-conditioned the subject's local impression of C_B. Away from C_B rear image expansion[1] is again evident, peaking at about 9° at L_B and R_B, with s.d.s of $\pm 4{\cdot}5^\circ$.

The percussive music programme excerpt used had considerable high-frequency spectral content, the first 15 seconds being mainly above 2 kHz and the latter 15 seconds mainly above 700 Hz. A programme excerpt was then

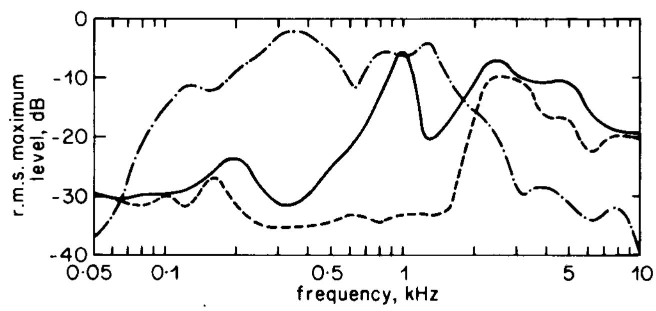


Fig. 7 - Spectral content of programme material used for subjective tests (as presented in the free-field room experiments)

——————— 30-second percussive music excerpt
- - - - - - - - first 15 seconds of percussive music excerpt
—·—·— 30-second male voice excerpt

selected having a greater low frequency content and the experiments were repeated. This consisted of a 30 second news excerpt read by a trained male announcer, and had a spectrum essentially confined below 2 kHz (see Fig. 7). Results for this material (Fig. 6 (in red)) are similar to those to those using percussive music.

It is concluded that 'surround-sound' reproduction systems should be capable of accurately reproducing the absolute azimuths of sounds in the front semi-circle and at C_B, but some degree of latitude ($\pm 10^\circ$) is permissible in the rear semi-circle near L_B and R_B.

4. Relative perception of direction

Undoubtedly a more stringent requirement of 'surround-sound' systems is their ability to differentiate between two closely spaced sound sources, since under these conditions the human auditory response is involved in making a comparison. The experiments described in the previous section were repeated in the free-field room using a second loudspeaker as a reference sound-source, and the subject was asked to move the test sound-source so as to be directly in line with the reference. In this experiment the sources were a matched pair of compact loudspeaker units, the test-source traversing just above the reference, with their high frequency units placed adjacent to one another so as to minimise the subjective height difference (see Fig. 8). The subject selected the reference or test loudspeaker unit by means of a switch.

Results for both types of programme material are shown in Fig. 9, and it is seen that the azimuthal acuity is much more accurate than in the case of a single source, and now greatest uncertainty occurs at the sides of the subject (s.d.s $\pm 3{\cdot}5^\circ$) where audition becomes largely monaural. Localisation errors are not significant at any azimuth and there are no significant differences between the results obtained with the two programme excerpts.

Relative azimuthal acuity is thus very accurate, and this is clearly an important factor in 'surround-sound' reproduction. The listener may not be aware of true positional errors of various sound sources, but is likely to be much more critical of their relative positions.

It is also of interest to note that during the experiment a number of subjects experienced front/back ambiguity on several occasions. When both loudspeakers were approximately at C_B the subject sometimes perceived one or both to be in mirror-image locations near C_F, and even when informed of their error sometimes had great difficulty in perceiving the true locations. It would therefore appear that there is auditory ambiguity on the front/back centre-line through the head, which would normally be resolved by head movement[6] or visual cues.

5. Azimuthal image perception

Quadraphonic systems rely on an extension of the well-known stereophonic image principle,[7] which provides

Fig. 8 - Free-field arrangement for relative sound localisation experiment

a sound image located (normally) between two loudspeakers placed in front of the listener, depending on the inter-channel level-difference.[7,8,9] By use of four loudspeakers placed symmetrically around the listener an extension of this principle might be expected to provide complete azimuthal coverage, although this requires that the angle subtended at the listener between adjacent loudspeakers be 90°, rather than the 60° more usually preferred in stereophony.

Accordingly an experiment was devised to determine the 'interchannel level-difference law' of image localisation for adjacent loudspeakers placed in a quadraphonic array. The generally preferred 'square array' (see Terminology) was used although qualitative comments on the 'diamond array' are noted below. A similar arrangement to that shown in Fig. 8 was employed, with four matched high-quality loudspeakers placed on a 2·7 m radius circle at positions L_F, R_F, R_B and L_B. A small locating loudspeaker was used to provide the moveable reference sound-source and was arranged so that it did not significantly disturb the sound field generated by the loudspeakers in the quadraphonic array. Absolute loudspeaker levels were adjusted, initially, to be equal at the observer's head location, using a sound-level meter, and then the relative levels of an adjacent pair was adjusted, maintaining the total power delivered to the two loudspeakers constant (cf. Appendix), until the subject judged that the image created was azimuthally coincident with the reference source. Noise masking was again used between tests whilst the reference loudspeaker was moved, and the subject had a single C_F reference marker to look at.

Initial tests rapidly showed that the free field room was an unsuitable environment in which to form subjective sound images from such pairs of loudspeakers, particularly at the sides of the head, because of the lack of a reverberant sound field, and so a wooden floor was constructed to provide a single, uniform, reflecting surface. This improved the localisation of images considerably, although loudspeaker location with respect to the listener was found to be very critical. A 2% misplacement in radial distance of one

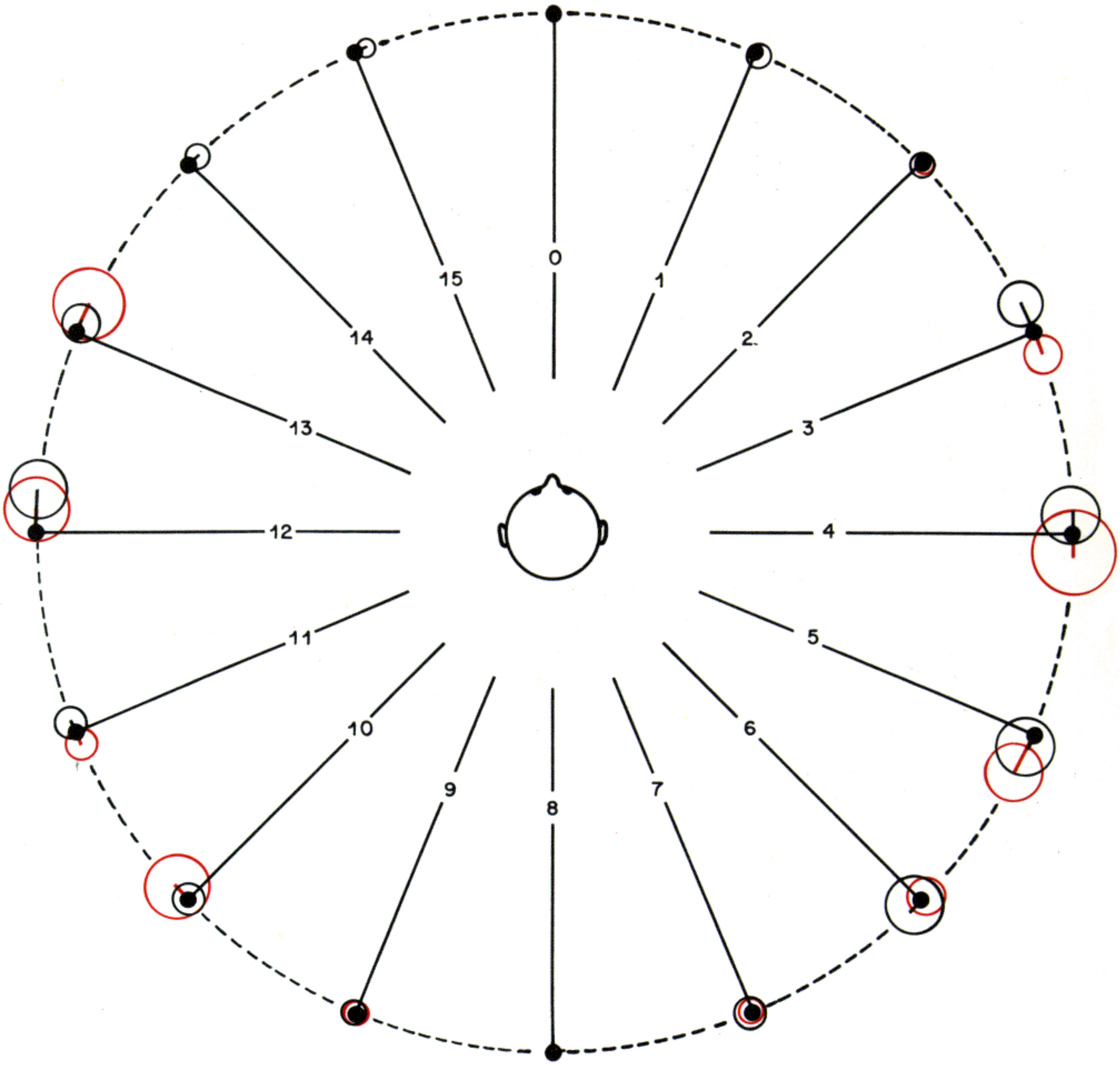


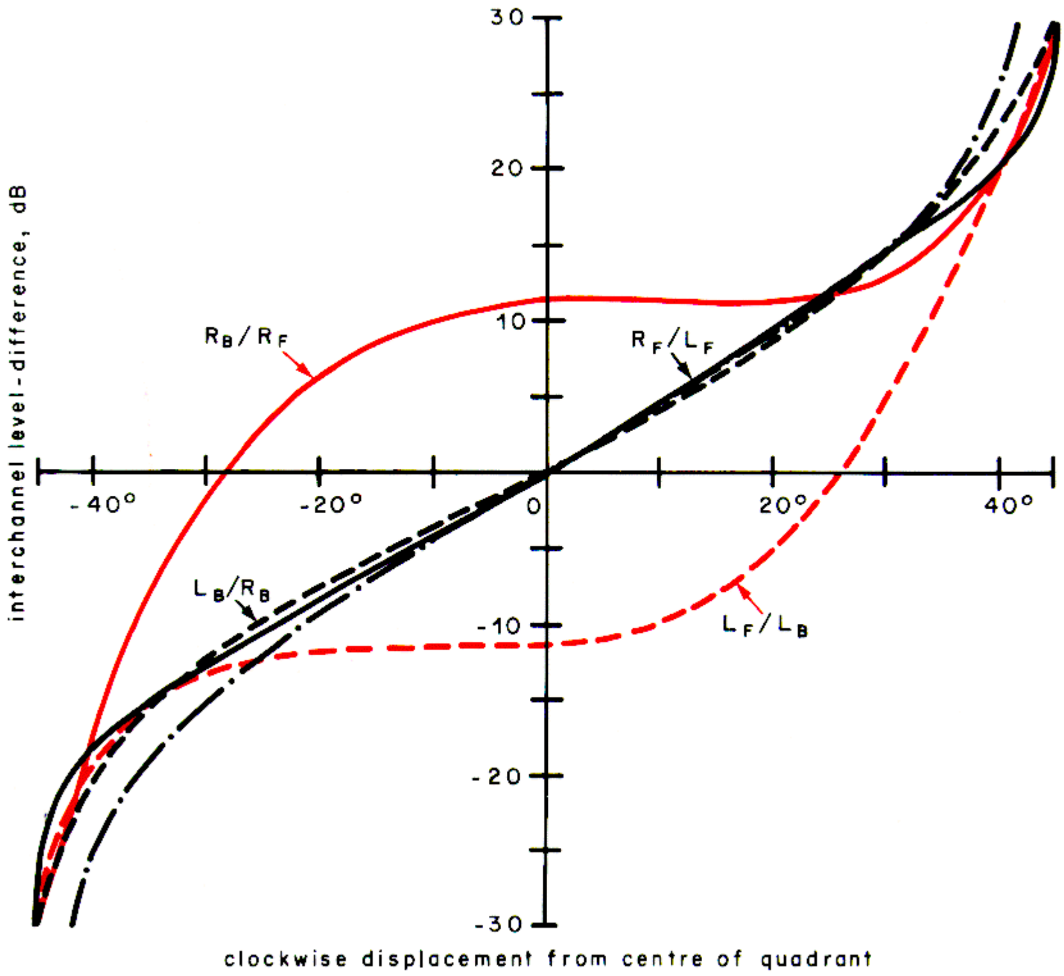
Fig. 9 - Relative sound localisation in the free-field room

reference source position

Percussive music results

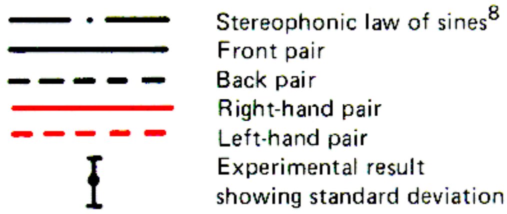
Male voice results

corresponding mean movable source position showing standard deviation circles

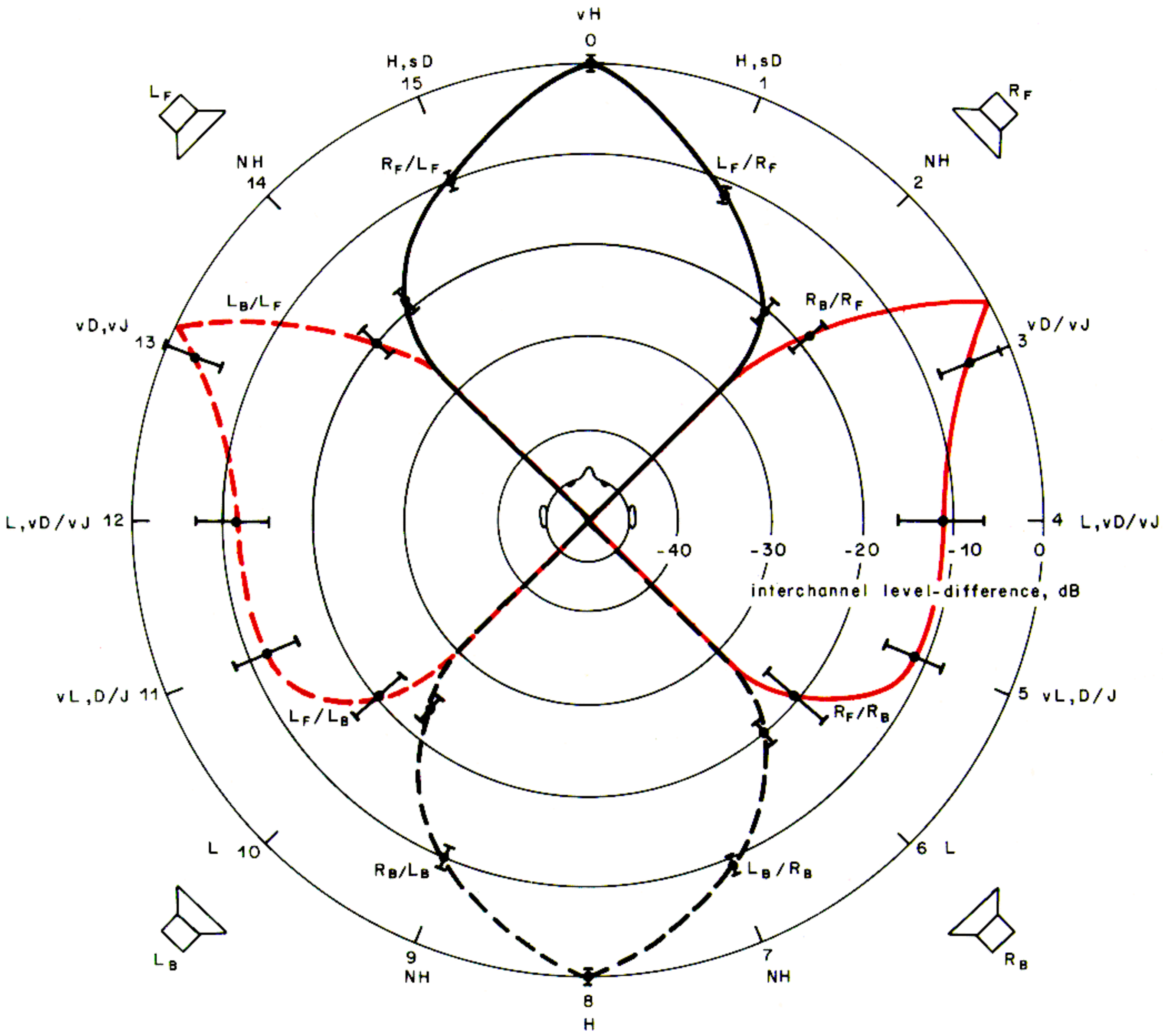


(*a*) Cartesian plot

Fig. 10 - Interchannel level-difference versus image location for adjacent pairs of loudspeakers in a quadraphone ('square') array, in the free-field room with a reflecting floor



D diffuse
H high
J jumpy
L low
NH normal height
s slightly
v very



(*b*) Polar plot

Fig. 10 - Interchannel level-difference versus image location for adjacent pairs of loudspeakers in a quadraphonic ('square') array, in the free-field room with a reflecting floor

Stereophonic law of sines[8]		D	diffuse
Front pair		H	high
Back pair		J	jumpy
Right-hand pair		L	low
Left-hand pair		NH	normal height
Experimental result showing standard deviation		s	slightly
		v	very

of the front loudspeakers (L_F and R_F) displaced a normal C_F image such that a 3 dB interchannel level-difference was required to correct it, whereupon the image exhibited a marked 'phasey' quality.

Results, the means obtained using seven subjects, are shown in Fig. 10 in both polar and rectilinear form. However, caution should be exercised in their interpretation, since large s.d.s were obtained for some locations, along, with adverse subjective comments (see polar plot). The front and rear quadrants are well defined and behave in the expected manner (cf. stereophony results[7,8,9]). However, the side quadrants exhibit a great degree of uncertainty, and subjects complained of either very diffuse or jumping images, with and without small head movements. It would appear that the 'stereophonic-image' phenomenon breaks down at the side of the head when predominantly one ear is excited, and the subject tends to hear each loudspeaker independently. Also there is preferential reception of the front loudspeaker, and about 10 dB more signal is required in the rear loudspeaker to give any impression of a centre-side image. There is a distinct threshold interchannel level-difference in this region, about which small variations cause the image to jump towards the front or back. However, the actual relative level at which this occurs varies greatly from subject to subject and from one occasion to another, as indicated by the large s.d.s for side image locations.

However, the effect is not so noticeable in the listening room, presumably because reflections cause directional information to be presented to both ears. It appears that the law determining image position is then largely dependent upon the phsyical properties of the room, and thus can be infinitely variable. For this reason, and lack of a totally symmetrical listening room, all the experiments on two-loudspeaker image-localisation were conducted in the free-field room with a single reflecting surface (i.e. a floor). Although such an environment does not give the same subjective impression as that of a listening room it does provide a far more critical and repeatable environment in which to determine the characteristics of various quadraphonic simulations; further, it gives good indications of possible short-comings, which may be subjectively disturbing under typical listening conditions.

Another feature of the image created by an adjacent pair of loudspeakers is the variation in image height around the compass. At C_F it is elevated by about 40° ('very high' in Fig. 10) and drops to eye-level height ('normal') at the front loudspeakers. Around the sides the image drops further becoming depressed by about 30° ('very low') at positions 5 and 11, such that the image appears to be at, or slightly below, floor level. Further towards the rear the image rises slightly, still being depressed by about 15° ('low') at the rear loudspeaker locations, and only rises slightly above eye-level (about 10° elevation) at C_B. The elevated front images are a seriously noticeable defect in quadraphony, although not serious in stereophony, and appear to be due to the increased angle subtended at the listener by the front loudspeakers. A theoretical hypothesis has been put forward to explain this effect,[8] but it does not agree well in magnitude, nor is there any reason to suggest an image rise as opposed to a fall.

Since side image localisation is poor with two-loudspeaker image synthesis it was considered possible that the 'diamond array' might prove more satisfactory. This configuration was briefly tested, and although image localisation was generally more similar in each quadrant, severe front/back ambiguity occurred as mirror-imaging about the C_L/C_R line. Accordingly this array was not considered further.

6. Effect of unwanted signals

So far only two-loudspeaker excitation has been considered in forming an image from a quadraphonic array, but in many quadraphonic systems three or even all four loudspeakers may be excited for a single point-source. In recording, for example, the use of four coincident cardioid microphones introduces unwanted* components (crosstalk) into a discrete quadraphonic reproduction, such that not only are the two adjacent channels to the source position energised, but also the two opposite channels carry components some 10–15 dB down.

Further experiments were conducted to give indications of the effects produced by exciting more than two loudspeakers in the array, and subjects were asked to determine the minimum perceptible crosstalk level for a number of selected situations, and to comment on localisation and quality changes brought about by excess crosstalk. Test image positions were chosen at either a loudspeaker (corner) location or mid-way between these (a centre-quadrant position), such that the wanted signal was applied either solely to one loudspeaker or equally to an adjacent pair. However, left/right symmetry was assumed and not all permutations were examined. Crosstalk signals were introduced into the diagonally opposite or adjacent pair of loudspeakers and the seven arrangements tested are shown in Fig. 11. The subject performed under the same test conditions as in the previous experiment, but was provided with a switch to add in the crosstalk, and an attenuator with which to vary its level. Having determined the minimum perceptible crosstalk level the subject was then asked to increase its level until it became equal to that of the wanted signals, describing the locus of the sound image as the increase was made. These results, the averages obtained using seven subjects, are shown in Fig. 11, and it is notable that in general about −20 dB of crosstalk is detectable, although it is considerably less in the C_F/C_B directions. As the crosstalk level is increased the images move in closer towards the subject, becoming bass heavy, and ending up rather unpleasantly within, or just above the subject's head.

* The signals from an adjacent pair of loudspeakers forming an image localised according to the 'interchannel intensity-difference' law (determined in the previous section) are termed the 'wanted' components and any radiated by the other loudspeakers are termed 'unwanted' or 'crosstalk' components. However, more generally, images formed by excitation of more than two loudspeakers are not necessarily undesirable, for instance, in the production of ranging effects.

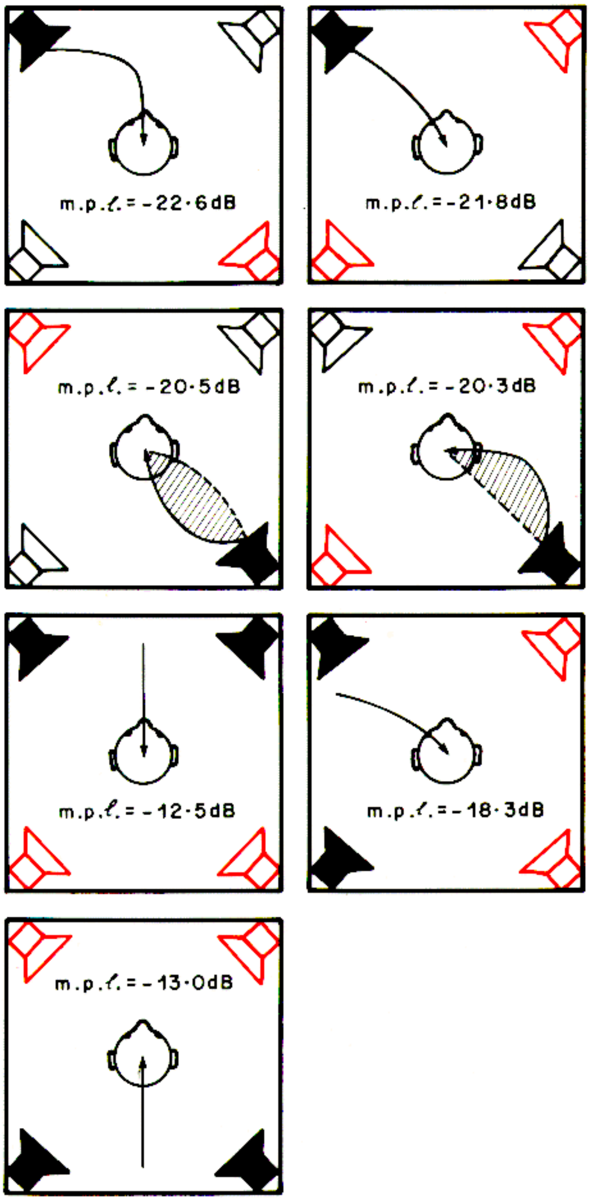


Fig. 11 - Minimum perceptible crosstalk level and image locus with increasing crosstalk for selected quadraphonic arrangements

m.p.*l.* is the minimum perceptible level of unwanted signal

wanted signal at reference level (0 dB)

unwanted (crosstalk) signal

no signal

locus of image position as crosstalk signal is increased from m.p.*l.* to 0 dB also showing area of uncertainty

7. Effect of phase differences

7.1. Occurrence

A number of quadraphonic systems employ only two transmission channels and matrix the four primary signals in differing amplitudes and phases into two composite signals. Decoding on reception produces crosstalk components with various amplitude and phase relationships relative to the wanted components, which themselves may exhibit phase differences. This section deals with a number of experiments devised to give some indication of the subjective effects experienced when such phase-shifted signals are presented.

7.2. Phase-shift between the wanted signals

In this experiment the minimum perceptible phase difference between the signals feeding an adjacent pair of loudspeakers was determined for selected image positions. The subject was tested under the same conditions as in the earlier image-locating tests, and was given a switch to compare the image with and without the inserted phase difference. The latter was reduced in 22½° steps (plus a final step to 11¼°) until the subject judged it only just perceptible. Seven image positions were tested, nominally at the centre of each quadrant and ±5° from each loudspeaker location; however, not all possibilities were tested since left/right symmetry was assumed. Fig. 12 shows results averaged from six subjects indicating the minimum perceptible phase difference for the selected image positions, and the corresponding azimuthal image-shifts which occurred. Centre-quadrant locations are most sensitive to phase differences, which is to be expected, but the side quadrant is considerably less sensitive than either front or back quadrants. Image shift follows the Haas precedence effect[10] at C_F and C_B, but when one loudspeaker is dominant the image always shifts towards it as the phase difference increases. Also at C_R (nominal position, equal signals to R_F and R_B loudspeakers) the image always moves forwards towards the front loudspeaker, presumably because the decorrelating effect of introducing phase-shift further enhances the forward source. General comments on the effects of excess phase-shift are that the images tend to move across the stage in a large arc, apparently moving further away from the observer as the phase difference increases, and becoming 'nasal' or bass lacking in quality. However, for images in a centre-quadrant location further increase in phase shift ($>90°$) causes the image to become diffuse, and finally results in the familiar unpleasant 'in the head' or 'phasey' sensations usually associated with stereophonic systems in which one loudspeaker has been phase-reversed.

7.3. Phase shift in the unwanted signals

The number of possible arrangements which could have been investigated is almost infinite, and so a very restricted set of tests was performed, based on the kinds of crosstalk components commonly introduced by existing matrix quadraphonic systems, and the tests previously reported in Section 6. The wanted signals were maintained in-phase and the unwanted signals were varied in 90° steps

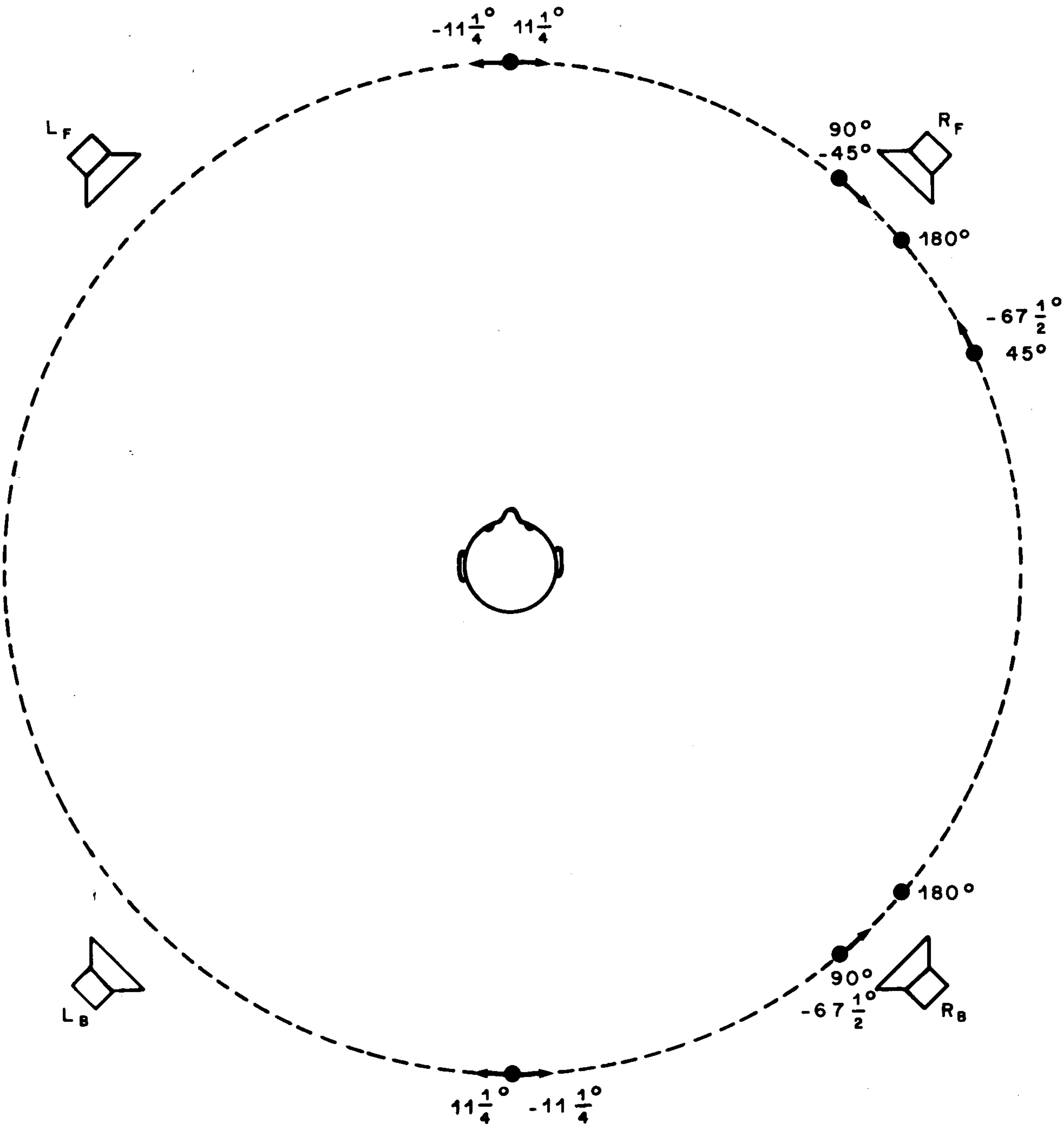


Fig. 12 - Minimum perceptible phase-difference between an adjacent pair of loudspeaker signals in a quadraphonic ('square') array showing image movement for selected positions

image location for in-phase signals

image location for minimum perceptible phase-shift

The sign of the phase-shift inserted is defined positive for the phase-leading loudspeaker clockwise of the nominal image position

(plus a minimum step of 45°), at a fixed level 10 dB below the wanted signals. The same seven loudspeaker configurations of Fig. 11 were used, and also an asymmetric crosstalk condition typical of the 'SQ' type of matrix was tested.

The subject was given a 3-way comparison of the wanted-signal image, and the composite-signal* image both with and without phase-shift applied to the crosstalk signals. This enabled him to identify readily the effects of the phase-shifts inserted. Results, the averages obtained with seven subjects, are presented in Fig. 13, and merit some explanation. Fig. 13(*a*) shows the effects of phase shift when the crosstalk signals are applied diagonally opposite the wanted signals, and Fig. 13(*b*) shows the results for corner images when two crosstalk signals occur in combinations similar to the 'QS' and 'SQ' types of matrix. Table I shows which loudspeakers were

* Wanted plus crosstalk signals.

Test Condition	Nominal Image Position	Crosstalk type	Relative loudspeaker levels (dB)			
			L_F	R_F	R_B	L_B
A	L_F	Symmetrical	0	–	–10	–
B	C_F	Symmetrical	0	0	–10	–10
C	C_R	Symmetrical	–10	0	0	–10
D	R_B	Symmetrical	–10	–	0	–
E	C_B	Symmetrical	–10	–10	0	0
F	L_F	'QS'	0	–10	–	–10
G	R_F	'SQ'	–	0	–10	–10
H	R_B	'QS'	–	–10	0	–10
I	L_B	'SQ'	–10	–10	–	0

Table I Test Conditions Referenced in Fig. 13

energised for each test condition (A to I) illustrated in Fig. 13(*a*) and (*b*). Crosstalk phase-shift was inserted both leading and lagging the wanted signals, and image position and quality comments are shown on concentric circles corresponding to a particular phase-shift. In cases where two crosstalk signals were present, phase-shift could be applied equally to both signals such that they were always in-phase, or alternatively with opposite sign such that one signal led, and one signal lagged, the wanted component. When both crosstalk signals were in-phase, or when only one crosstalk signal was present, the sign of the phase-shift was not subjectively detectable, and is therefore not indicated in the figure. However, when the crosstalk signals were phase-shifted in opposite directions from the wanted signals, a sign is appended to the phase-shift indicated in the figure; this is considered to be positive when the crosstalk signal located in the clockwise direction relative to the image position is phase-leading.

Generally, in-phase crosstalk produces images closer to the observer and bass heavy in quality (as was experienced in the earlier crosstalk tests), whereas anti-phase crosstalk produces phasey or nasal, bass lacking images. In the side quadrants in-phase crosstalk produces image shifts towards the front/back centre-line, and anti-phase crosstalk towards the left/right centre-line. With two crosstalks present simultaneously, in-phase crosstalk signals produce no image shift, but the image generally sounds phasey and diffuse. However, if one crosstalk signal leads and the other lags the wanted signals, less objectionable image qualities are observed, although some azimuth movement may be observed. Closer inspection of the results shows that, in general, crosstalk phase shifts of 45°, or +45° and –45° for two signals, produces the least disturbing subjective effect, except at centre-side where +90° and –90° is preferred.

Referring to Fig. 13(*c*) observation of the 'QS type' corner crosstalk (test conditions F and H) shows distinctly asymmetric results, although the exact locus of image movements is not well defined owing to the relatively small number of tests conducted. However, +45° and –45° again appear to give a satisfactory image quality with little image shift. With 'SQ type' crosstalk (test conditions G and I) the signal opposite the wanted signal was always in-phase or in anti-phase, and the adjacent signal phase is either + or –90° as indicated by the '0°/+90°' and '180°/–90°' nomenclature in the figure. Large image shifts or poor image quality is experienced at the front corners, the shift being dependent on the sign of the adjacent signal (Haas effect applied). At the rear corners the image shift is very small and quality not so impaired. However, in both cases anti-phase diagonal crosstalk is preferable to the in-phase form.

8. Comment on the results

These experiments have explored some of the fundamental properties of hearing which have particular bearing upon the engineering of 'surround sound' reproduction systems. Also, the effects of four loudspeaker or quadraphonic presentation of sounds have been investigated, and some insight into such limited point-source simulations of sound-fields has been gained.

Fundamental properties of hearing determined are:

(a) the human auditory system is approximately of equal sensitivity to isolated sounds from all azimuths (see Fig. 2);

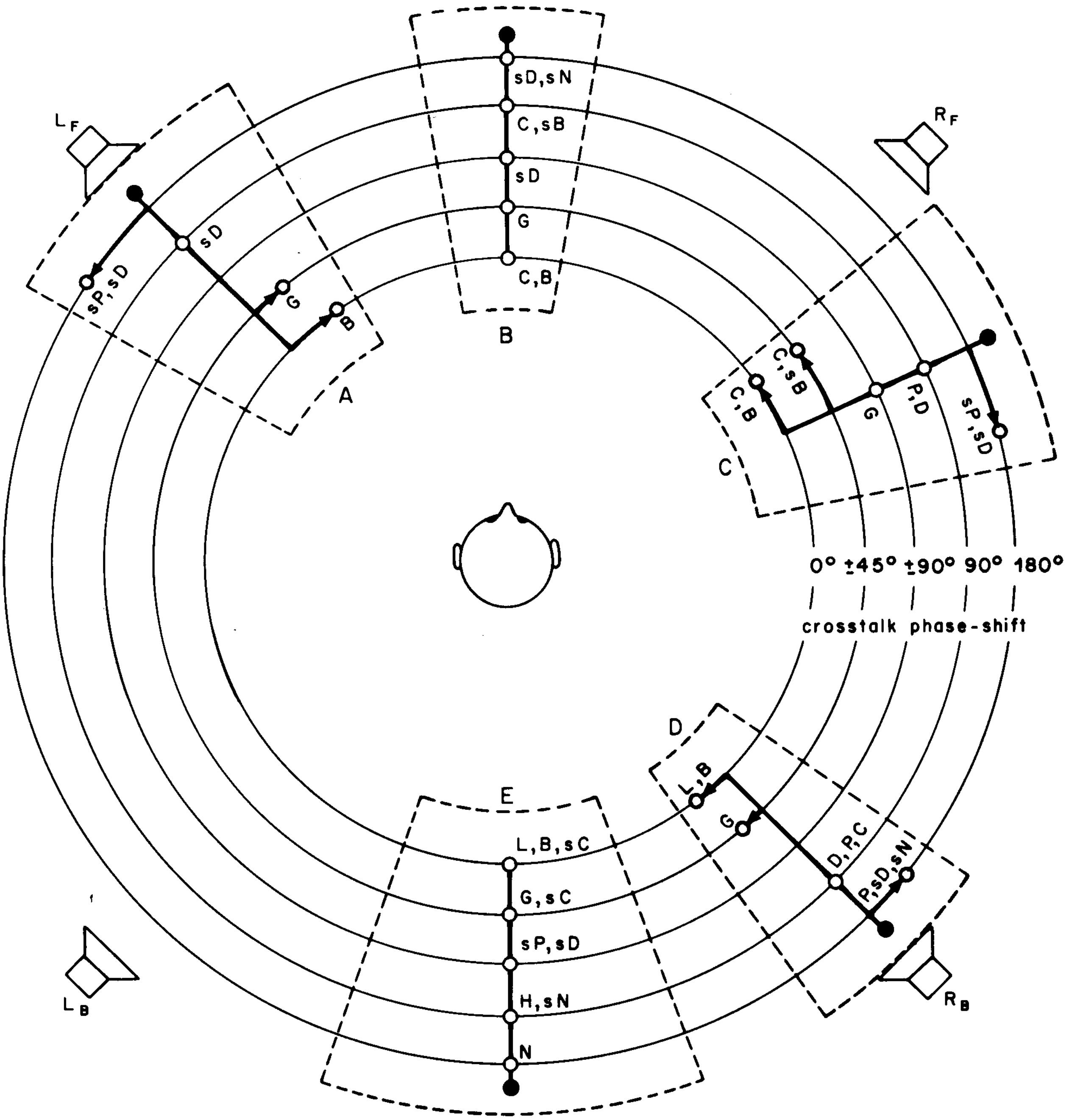


Fig. 13(a) - Effect of phase-shifted crosstalk signals on image location and image quality for selected quadraphonic arrangements:
Crosstalk image diagonally opposite wanted image position (test conditions A to E shown in Table I)

image location without crosstalk

image locations with crosstalk showing shift (the magnitudes of the crosstalk phase-shifts are indicated by the concentric circles)

B	bass heavy	H	high) vertical position
C	close, near	L	low) vertical position
D	diffuse	N	nasal, bass lacking
F	far, distant	P	phasey
G	good, no adverse comments	s	slightly

The phase-shift inserted into the crosstalk (unwanted) signals is defined positive for the phase-leading crosstalk clockwise of the nominal image position, and the equally phase-lagging signal anticlockwise. No sign indicates that all crosstalk signals are in-phase, but either lead or lag the wanted signals by the specified amount.

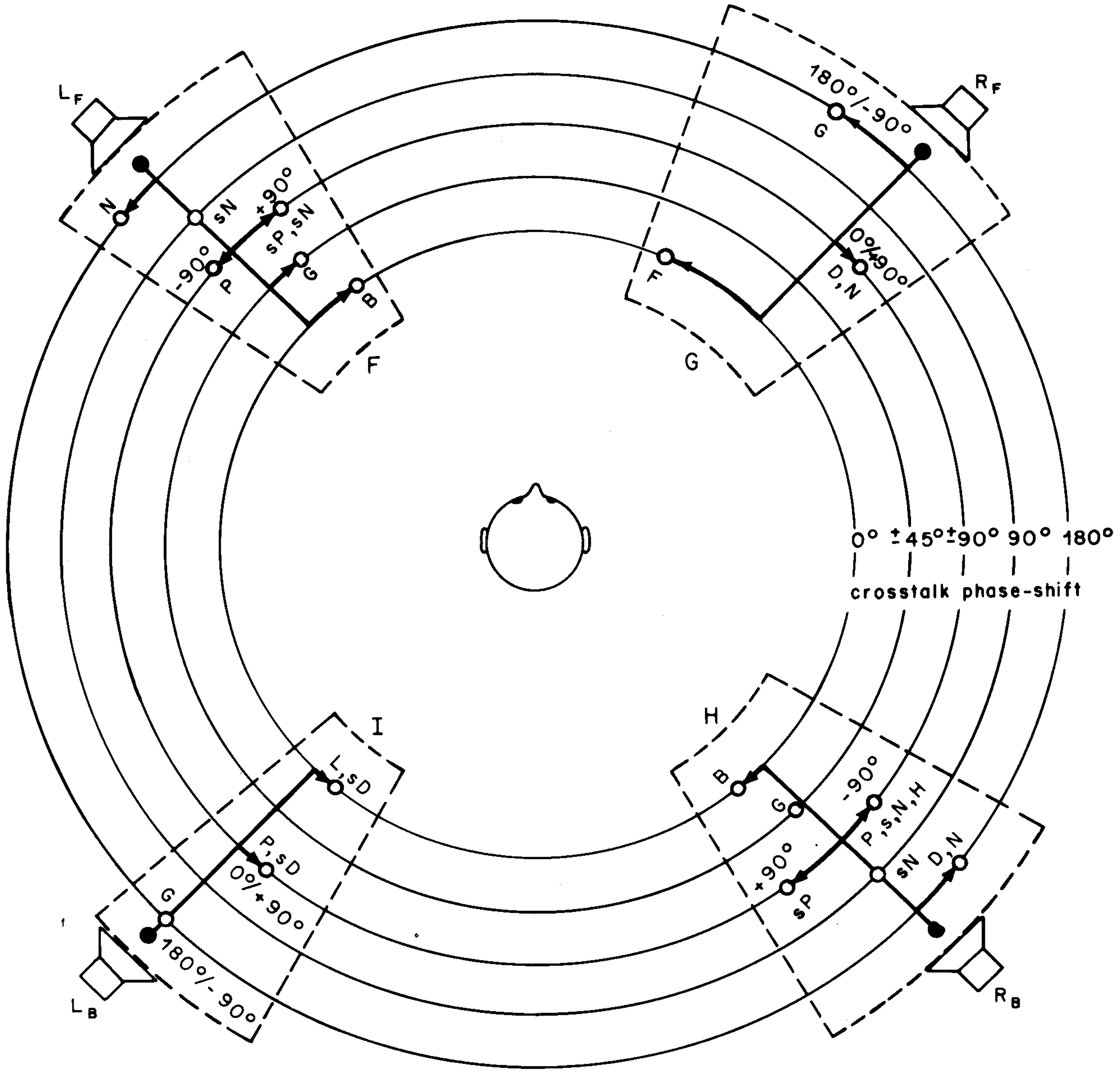


Fig. 13(b) - Effect of phase-shifted crosstalk signals on image location and image quality for selected quadraphonic arrangements:
Crosstalk images of the 'SQ' and 'QS' types at corner positions (test conditions F to I shown in Table I)

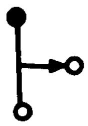


image location without crosstalk

image locations with crosstalk showing shift (the magnitudes of the crosstalk phase-shifts are indicated by the concentric circles)

B	bass heavy	H	high) vertical position
C	close, near	L	low)
D	diffuse	N	nasal, bass lacking
F	far, distant	P	phasey
G	good, no adverse comments	s	slightly

The phase-shift inserted into the crosstalk (unwanted) signals is defined positive for the phase-leading crosstalk clockwise of the nominal image position, and the equally phase-lagging signal anticlockwise. No sign indicates that all crosstalk signals are in-phase, but either lead or lag the wanted signal by the specified amount.

(b) isolated sound-source localisation in the horizontal plane is accurate to about 5° in the front semi-circle and at centre-back (C_B), but greater uncertainty (≏ 10°) exists, and considerable rear-image expansion occurs, in the rear semi-circle away from C_B (see Figs. 5 and 6);

(c) relative sound-source localisation in the horizontal plane is accurate to within 2° in front and rear quadrants, becoming more uncertain towards the centres of the side quadrants (≏ 5°) where audition is largely monaural (see Fig. 9).

The extent to which a quadraphonic 'square' array can reproduce a desired sound stage was investigated on the basis of an extension of stereophonic principles, and an interchannel level-difference law of image position for adjacent pairs of sources has been determined around the complete compass (see Fig. 10). However, the ability of the listener to realise an image so formed at the sides of the head was found to be questionable under non-reverberant conditions, and considerable front-source dominance* and independent reception of the front and rear sources was then evident. It was also found that room acoustics play an important part in specifying the interchannel level-difference law in these regions.

Fig. 10 shows the fidelity of back image localisation according to normal stereophonic principles,[8] which is not in agreement with the 'back image contraction' principle[11] claimed by workers elsewhere. However, the effect cited in favour of the latter principle may be explained by the apparent expansion of the real stage when the observer turns his back on it (see Figs. 5 and 6).

Perceptibility of unwanted (crosstalk) signals (Fig. 11) is high (at a relative level of about −20 dB) when one source is dominant (i.e. near corner locations), but is reduced (to a relative level of about −12 dB) for images in the centre-front and centre-back areas. Excess crosstalk produces images close to the subject and bass heavy in quality.

This conflicts with the 'front source dominance principle'[11] and indicates that the strongest source always predominantly defines the location of the image. Only with adjacent channel crosstalk will the image tend to lie to the forward side of the dominant loudspeaker. A diagonal crosstalk signal (coming from the loudspeaker opposite the dominant one) however, will always cause an image shift towards the front/back centre-line.

Investigation of the permissible phase-shift between adjacent pairs of loudspeakers (cf. results of Ref. 12) has shown greatest sensitivity at centre-front and centre-back (≏ 11°), with image shifts occurring according to the Haas precedence effect (see also 'quadrature image-shift principle'[11]). Elsewhere, however, the image moves towards the dominant source or, in cases where a side image is produced, front-source dominance already experienced with no phase-shift is further enhanced.

If crosstalk components are introduced, phase-shifted with respect to the wanted signals, the resulting image may exhibit a wide range of subjective effects, and the results of the restricted set of tests conducted show that certain amounts of phase shift, applied to the crosstalk signals in specified senses, can improve the image quality and cause minimal shift from the desired image position. A good compromise between the bass heaviness produced by in-phase crosstalk signals and the 'phaseyness' produced by those in anti-phase with the wanted signals can be obtained, generally by applying about 45° phase-shift to a single crosstalk signal, or about +45° and −45° phase-shift to a pair. However, the combination of +90° and −90° is preferred for centre-side images, whereupon the listener's ability to realise the image is greatly enhanced.

It is thought preferable to arrange the sense of paired phase-shifted crosstalk components such that the Haas effect aids image localisation towards the left/right centre-line. This will tend to counter natural tendencies to localise the image towards the front/back centre-line, although the effect is often not significant.

9. Conclusions

Although of limited nature, these results illuminate the fallacy of generalisation from a few observed phenomena,[11] and confirm that the mechanism of audition is indeed extremely complex. Accordingly, attempts to deceive the normal auditory processes (i.e. the re-creation of a total sound stage by multiple (4) point-source simulation) should be extensively studied and understood before the optimum engineering solution to the problem can be found.

Many workers have been searching for effective methods of reducing four or more studio signals into a smaller number of transmission channels (typically two) by simple linear processing techniques (matrixing), such that the received signals may be further processed to reproduce the original omnidirectional sound stage. Such signal processing typically provides quadraphonic signals of the type investigated in this report, and the results presented provide pointers for the design of more effective systems, and have been used to pin-point the causes of the undesirable features of some systems presently in existence.[13] In this way two-channel matrix systems have already been evolved which provide, in some ways, subjectively more satisfying results than current commercial matrix systems. Work continues in this field.

Also of importance is the subjective susceptibility to the geometry of the loudspeaker array. A 2% misplacement of one loudspeaker caused a considerable image shift from a nominally centre-front location, and it was also found that image localisation was very sensitive to head position during many of the tests. For 'surround-sound' systems to be of much value in a domestic listening environment, reasonable tolerances on the geometry of the loudspeaker

* Not to be confused with the 'front source dominance principle' of Reference 11, which states that the 'human hearing mechanism will judge the direction of sound arrival based upon the signals proceeding from the front loudspeakers' providing that one of these signals is considerably greater than any of the others.

array and the permissible listening area must be allowed, and it is therefore considered necessary to study these aspects further.

10. References

1. GARDNER, M.B. 1973. Some single and multiple source localisation effects. *J. Audio Engng Soc.*, 1973, **21**, 6, pp. 430 – 437.

2. FRANSSEN, N.V. 1964. Stereophony. Eindhoven, Philips Tech. Lib., 1964.

3. TOBIAS, J.V. 1971. Foundations of modern auditory theory. Vol. II, New York, Academic Press, 1972.

4. DE BOER, K. 1940. Stereophonic sound reporduction. *Philips Tech. Rev.*, 1940, **5**, 4, pp. 107 – 114.

5. SHAXBY, J.H. *and* GAGE, F.H. 1932. The localisation of sounds in the median plane. Med. Red. Council Spec. Rept. Series No. 166, London, HMSO, 1932.

6. WALLACH, H. 1940. The role of head movements and vestibular and visual cues in sound localisation. *J. Exp. Psych.*, 1940, **27**, 4, pp. 339 – 368.

7. CLARK, H.A.M., DUTTON, G.F. *and* VANDERLYN, P.B. 1957. The 'Stereosonic' recording and reproducing system. *Proc. Instn elect. Engrs.*, 1957, **104B**, 17, pp. 417 – 432.

8. BAUER, B.B. 1961. Phasor analysis of some stereophonic phenomena. *J. Acoust. Soc. Am.*, 1961, **33**, 11, pp. 1536 – 1539.

9. HARWOOD, H.D. 1968. Stereophonic image sharpness. *Wireless Wld.*, 1968, **1393**, 74, pp. 207 – 211.

10. HAAS, H. 1951. The influence of a single echo on the audibility of speech. (trans. from Acustica, 1951, **1**, pp. 49 – 58), *J. Audio Engng Soc.*, 1972, **20**, 2, pp. 145 – 159.

11. BAUER, B.B., GRAVEREAUX, D.W. *and* GUST, A.J. 1971. A compatible stereo-quadraphonic (SQ) record system. *J. Audio Engng Soc.*, 1971, **19**, 8, pp. 638 – 646.

12. KOHSAKA, O., SATOH, E. *and* NAKAYAMA, T. 1972. Sound-image localisation in multi-channel matrix reproduction. *J. Audio Engng Soc.*, 1972, **20**, 7, pp. 542 – 547.

13. CROMPTON, T.W.J. The subjective performance of various quadraphonic matrix systems. BBC Research Department Report in course of preparation.

Appendix

Equal Loudness Levels of a Spaced Pair of Loudspeakers and a Single Loudspeaker

A preliminary investigation was conducted into the equal loudness levels for an image formed by a pair of spaced loudspeakers and a single loudspeaker placed at the image location. A pair of loudspeakers was set up in the listening room, 3 m apart, such that they subtended an angle of 90° at the listener. A similar single loudspeaker was placed on the bisector of this angle at the same radial distance (1·7 m) from the listener, and the latter faced this loudspeaker. The listener was asked, by making a switched comparison, to adjust the level of the signal fed to the loudspeaker-pair such that the perceived loudness of this image matched that of the single loudspeaker. White and pink noise test signals were used, and the signal level to the single loudspeaker was fixed to give a sound pressure level (s.p.l.) about 65 dBA at the listener's head position.

Results averaged from six subjects show an attenuation of 5·2 dB (s.d. = ±1·2 dB) on the signal feeding the loudspeaker-pair over that feeding the single loudspeaker when using white noise, and 4·5 dB (s.d. = ±1 dB) using pink noise. The measured s.p.l. difference at the listener's head was +5 dB for the loudspeaker-pair radiating at the same levels as the single loudspeaker using either white or pink noise.

For comparison the experiment was repeated with the loudspeaker-pair subtending a 60° angle at the listener, as in normal stereophony. The subjective level difference was reduced slightly to 4·1 dB (s.d. = 1·3 dB) using white noise, but was still 4·5 dB (s.d. = 1 dB) using pink noise. The measured s.p.l. difference was also still 5 dB using white or pink noise.

SMW/AMM